Conversion of communicated speech to text for tranmission as RF modulated base band video

Patent Number: US5774857
Publication date: 1998-06-30
Inventor(s): NEWLIN DOUGLAS J (US)
Applicant(s):: MOTOROLA INC (US)
Requested Patent: DE19750439
Application Number: US19960751048 19961115
Priority Number(s): US19960751048 19961115
IPC Classification: G10L3/00
EC Classification:
Equivalents: AU4443697, BR9705853, CN1190840, GB2319390

Abstract

Apparatuses, systems, and a method provide for a visual display of speech, such as the visual display of a received audio signal in telecommunications, especially useful for the hearing impaired. The preferred apparatus includes a network interface that is coupleable to a first communication channel to receive an audio signal; a radio frequency (RF) modulator to convert a baseband output video signal to a RF output video signal and to transmit the RF output video signal on a second communication channel for video display; and a processor coupled to the network interface and to the RF modulator for running a set of program instructions to convert the received audio signal to a text representation of speech, and to further convert the text to the baseband output video signal. The RF output video signal, when displayed on a video display, provides the visual display of speech. The preferred apparatus may also include a speech generation subsystem.

(19) BUNDESREPUBLIK DEUTSCHLAND

DEUTSCHES PATENTAMT

(12) Offenlegungsschrift

(10) DE 197 50 439 A 1

(51) Int. Cl.[6]:
A 61 F 11/04
G 10 L 7/08
G 09 B 21/00
H 04 M 1/00
H 04 N 5/445

(21) Aktenzeichen: 197 50 439.6
(22) Anmeldetag: 14. 11. 97
(43) Offenlegungstag: 20. 5. 98

DE 197 50 439 A 1

(30) Unionspriorität:
751048 15. 11. 96 US

(71) Anmelder:
Motorola, Inc., Schaumburg, Ill., US

(74) Vertreter:
Dr. L. Pfeifer und Kollegen, 65203 Wiesbaden

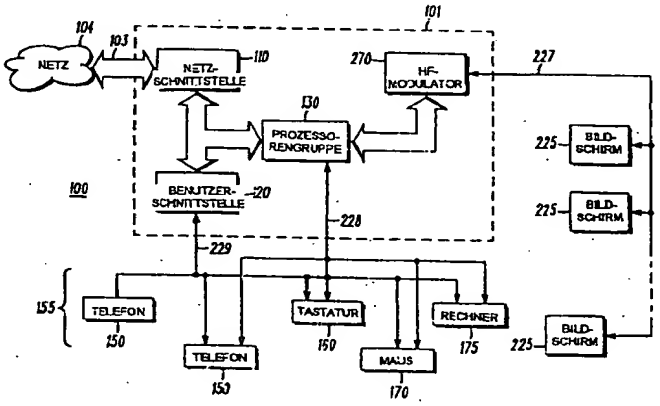
(72) Erfinder:
Newlin, Douglas J., Geneva, Ill., US

Die folgenden Angaben sind den vom Anmelder eingereichten Unterlagen entnommen

Prüfungsantrag gem. § 44 PatG ist gestellt

(54) Verfahren, Vorrichtung und System zur visuellen Wiedergabe von Sprache in der Sprachkommunikation

(57) Erfindungsgemäße Vorrichtungen (101, 201, 301), Verfahren und Systeme (100, 200, 300) sorgen für die visuelle Wiedergabe von Sprache, beispielsweise die visuelle Anzeige eines empfangenen Tonsignals bei der Telekommunikation, die besonders für Hörbehinderte nützlich sind. Das bevorzugte Ausführungsbeispiel der Vorrichtung umfaßt eine Netzschnittstelle (110), wobei die Netzschnittstelle mit einem ersten Kommunikationskanal zum Empfangen eines ersten Tonsignals zur Bildung eines Tonempfangssignals koppelbar ist, ferner einen Hochfrequenzmodulator (270) zur Umwandlung eines Bildausgangssignals im Basisband in ein Hochfrequenz-Bildausgangssignal und zur Übermittlung des Hochfrequenz-Bildausgangssignals auf einem zweiten Kommunikationskanal zur Bildwiedergabe, und eine mit der Netzschnittstelle und dem Hochfrequenzmodulator gekoppelte Prozessorengruppe (130), wobei die Prozessorengruppe (130) über einen Satz Programmbefehle in der Weise ansteuerbar ist, daß sie das Tonempfangssignal in eine Sprachwiedergabe in Textform umwandelt und außerdem die Textdarstellung gesprochener Sprache in ein Bildausgangssignal im Basisband umwandelt. Dabei bildet das Hochfrequenz-Bildausgangssignal bei Darstellung auf einem Bildschirm (225) die visuelle Wiedergabe der gesprochenen Sprache. Das bevorzugte Ausführungsbeispiel kann des weiteren ein Teilsystem zur Spracherzeugung umfassen.



DE 197 50 439 A 1

Beschreibung

Gebiet der Erfindung

Die vorliegende Erfindung bezieht sich ganz allgemein auf die Ton- und Bildkommunikation und insbesondere auf eine Vorrichtung, ein Verfahren und ein System zur visuellen Wiedergabe von Sprache in der Kommunikation.

Stand der Technik

Herkömmliche Geräte und Verfahren zur visuellen Wiedergabe gesprochener Sprache wie beispielsweise die sogenannten TDD-Systeme für Hörbehinderte bzw. Hörgeschädigte setzen im typischen Fall sowohl spezielle Systeme als auch eine Eingabe des anzuzeigenden Materials für die visuelle Wiedergabe durch den Benutzer voraus. Beispielsweise wird für Telephongespräche oder Kommunikationssitzungen für den Hörbehinderten ein spezielles TDD-System für die visuelle Anzeige von Buchstaben, Wörtern und Sätzen vorausgesetzt, und dabei müssen alle Teilnehmer an der Kommunikationsverbindung ein solches speziell hierfür vorgesehenes System benutzen. Außerdem muß jeder Teilnehmer bei dem Telefongespräch bei Verwendung eines TDD-Systems jeden Buchstaben, jedes Wort und jeden Satz physikalisch auf einer Tastatur eingeben, damit diese Informationen dann zur Anzeige auf einem TDD-System am entfernt liegenden Ende übermittelt werden.

Bei anderen konventionellen Systemen ist außerdem besonders ein Eingreifen von Hand erforderlich, wobei das anzuzeigende visuelle Material separat körperlich eingegeben werden muß. Beispielsweise ist es bei vielen Untertiteldiensten für geschlossene Benutzergruppen, wie sie auf vielen Fernsehkanälen zur Verfügung stehen, erforderlich, daß die hörbaren gesprochenen Wörter von einem Diensteanbieter übersetzt und zur Übertragung als Teil der Ton-/Bildsendung oder einer anderen Fernsehsendung in das Untertitelsystem für die geschlossene Benutzergruppe mittels Tastatur eingegeben werden.

Diese konventionellen Systeme zur visuellen Anzeige gesprochener Sprache setzen im allgemeinen zweckbestimmte Spezialsysteme sowohl vor Ort wie auch am entfernten Punkt der Verbindung voraus und erfordern einen erheblichen Umfang manueller Eingriffe für den Betrieb. Infolgedessen sind derartige Systeme relativ kostspielig und schwierig zu bedienen. Außerdem unterliegen Systeme dieser Art Beschränkungen hinsichtlich ihrer Verfügbarkeit und Aufstellung; beispielsweise können diese TDD-Systeme auf Reisen nur mit Schwierigkeiten aufgestellt oder lokalisiert werden, so daß die Kommunikation mit einer hörbehinderten Person über das Telefon unmöglich wird. Außerdem kann der Benutzer, der auf ein derartiges System angewiesen ist, nicht mit einem anderen Teilnehmer kommunizieren, dem ein spezielles System für diesen Zweck nicht zur Verfügung steht.

Dementsprechend blieb Bedarf an einem solchen Gerät, Verfahren und System zur visuellen Sprachanzeige, bei denen spezielle Geräte und Systeme nicht an beiden Enden der Kommunikationsverbindung vorhanden sein müssen. Außerdem sollten ein Gerät und ein System dieser Art keinen erheblichen Aufwand an manueller Betätigung für den Betrieb erfordern, sie sollten vergleichsweise kostengünstig und außerdem benutzerfreundlich sein.

Kurzbeschreibung der Zeichnung

Fig. 1 zeigt ein Blockschaltbild zur Darstellung eines erfindungsgemäßen Geräts und Systems zur visuellen Sprachwiedergabe;

Fig. 2 ist ein Blockschaltbild mit der Darstellung eines ersten bevorzugten Ausführungsbeispiels eines erfindungsgemäßen Geräts und Systems zur visuellen Sprachwiedergabe;

Fig. 3 zeigt ein Blockschaltbild zur Darstellung eines zweiten bevorzugten Ausführungsbeispiels eines erfindungsgemäßen Geräts und Systems zur visuellen Sprachwiedergabe; und

Fig. 4 ist ein Ablaufdiagramm zur Veranschaulichung eines erfindungsgemäßen Verfahrens zur visuellen Sprachwiedergabe und Spracherzeugung.

Ausführliche Beschreibung der Erfindung

Wie vorstehend bereits angesprochen blieben verschiedene Bedürfnisse für Möglichkeiten zur visuellen Sprachwiedergabe, beispielsweise unter anderem in einem Textformat oder einem Untertitelformat, als Hilfsmittel für Hörbehinderte bestehen. Die erfindungsgemäße Vorrichtung mit zugehörigem Verfahren und System baut auf den verwandten und damit zusammenhängenden Anmeldungen auf und sorgt für die visuelle Wiedergabe gesprochener Sprache, ohne daß hierfür vor Ort und am entfernten Ende der Kommunikationsverbindung hierzu spezielle Geräte und Systeme erforderlich sind. Außerdem setzen auch die verschiedenen Ausführungsbeispiele der Erfindung ebenfalls keinerlei erhebliche Betätigungseingriffe von Hand für den Betrieb voraus und sind dabei vergleichsweise kostengünstig und benutzerfreundlich.

Die in den verschiedenen hiermit zusammenhängenden Anmeldungen beschriebenen Erfindungen beziehen sich sowohl auf die Telefonkonferenztechnik als auch auf audiovisuelle Konferenztechnik und arbeiten mit einer Vorrichtung für den Zugriff auf Video- bzw. Bildinformationen, welche über einen Kommunikationskanal mit einem Telekommunikationsnetz gekoppelt werden kann. In der zweiten und dritten hiermit zusammenhängenden Anmeldung bezieht sich das dort bevorzugte Ausführungsbeispiel auf die Vorrichtung für den Zugriff auf Bildinformationen sowie für audiovisuelle Konferenztechnik unter Heranziehung eines sogenannten CACS-Protokolls (Cable ACcess System; Kabelzugangssystem) zur Kommunikation mit einer Hauptstation über ein koaxiales Hybrid-Koaxkabel, wobei die Primärstation ihrerseits für die Anschlußmöglichkeiten an ein Telekommunikationsnetz und eine Infrastruktur für Kabelfernsehdienste sorgt. Bei der hiermit zusammenhängenden vierten und fünften Anmeldung sieht die Vorrichtung für den Zugriff auf Bildinformationen sowohl Telekonferenzmöglichkeiten als auch Möglichkeiten für Audio-/Video-Konferenzen mit direkten festverdrahteten Anschlußmöglichkeiten an ein Telekommunikationsnetz vor, wobei eine festverkabelte Netzschnittstelle eingesetzt wird, die sich beispielsweise für den Anschluß an ein ISDN-Netz (Integrated Services Digital Network; digitales Netz mit integrierten Diensten) und/oder an ein PSTN-Netz (Public Switched Telephone Network; öffentliches Telefonnetz mit Wählsystem) eignet.

Bei den bevorzugten Ausführungsbeispielen der hiermit zusammenhängenden zweiten und vierten Anmeldung ist die Möglichkeit für Videokonferenzen unter Verwendung üblicher oder allgemein bekannter Geräte und Vorrichtungen vorgesehen, wie sie typischerweise in Räumen oder bei Teilnehmern zu finden sind, z. B. Telefone, Fernseher und Videokameras (Video-Camcorder). Bei der hiermit zusammenhängenden dritten und fünften Anmeldung ist eine solche Möglichkeit zur Videokonferenz unter Verwendung eines oder mehrerer Bildtelefongeräte vorgesehen. Was allerdings allen diesen hiermit zusammenhängenden Anmeldun-

gen gemeinsam ist, ist die Verwendung einer physikalischen Schnittstelle (z. B. in Form eines Telefons oder einer Tastatur) für die Auswahl und die Ansteuerung der verschiedenen Medieneinsatzgebiete, z. B. zur Auswahl eines normalen Telefonmodus oder eines Videokonferenzmodus. Bei den bevorzugten Ausführungsbeispielen können ein oder mehrere Telefone zur Eingabe verschiedener Steuersignale in eine Benutzertonschnittstelle des Videozugriffsgeräts eingesetzt werden, um die jeweilige Betriebsart des Geräts für den Zugriff auf Bildinformationen anzuwählen. Bei dem bevorzugten Ausführungsbeispiel wird beispielsweise mit der Eingabe einer vorgegebenen Abfolge (z. B. "**" der DTMF-Töne eines Telefons) gearbeitet, um einen Videokonferenzmodus anzuwählen, wobei dann, wenn diese vorgegebene Abfolge nicht eingegeben wurde, automatisch ein Telefonbetrieb in transparenter Weise gewählt wird.

Bei der hiermit zusammenhängenden sechsten Anmeldung sind eine Vorrichtung und ein Verfahren zur Ansteuerung mehrerer unterschiedlicher Multimedia-Anwendungen vorgesehen, neben Möglichkeiten für Videokonferenzen und Telefonbetrieb. Bei dem bevorzugten Ausführungsbeispiel der Erfindung gemäß dieser sechsten Anmeldung sorgt die Vorrichtung zur Multimedia-Ansteuerung für die Kontrolle über eine Vielzahl von Medienanwendungen, unter anderem Telefon, Videokonferenz, analoge und digitale Videotechnik sowie Signalabgabe über die Wechselstromleitungen (zur Ansteuerung und Überwachung von Geräten im Raum oder beim Teilnehmer, z. B. Heizung, Lüftung, Klimaanlage, Beleuchtung, Sicherheitseinrichtungen und Unterhaltungstechnik). Darüber hinaus kann bei dem bevorzugten Ausführungsbeispiel der Multimediasteuerung jedes angeschlossene Telefon zum Telefon für mehrere Betriebsarten werden, wobei es die physikalische Schnittstelle für Telefonfunktionen und für Multimedia-Steuerfunktionen bildet.

Auf diesen hiermit zusammenhängenden Anmeldungen bauen die erfindungsgemäße Vorrichtung, das Verfahren und das System auf und sehen eine visuelle Wiedergabe gesprochener Sprache vor, beispielsweise bei einem Sprachtelefongespräch oder dem Audioteil einer audiovisuellen Konferenz. Die Kommunikation kann über jedes Telekommunikationsnetz oder auch jedes andere Netz ablaufen, wobei am entfernten Punkt der Verbindung kein besonderes oder spezielles Gerät erforderlich ist. Wie im folgenden noch ausführlicher beschrieben ist, wird ein aus einem Netz ankommendes Tonsignal empfangen und in eine Darstellung in Textform umgewandelt, die dann in ein Bildsignal umgesetzt wird, das in jedes angeschlossene Fernsehgerät oder ein anderes Bildschirmgerät übertragen wird, wo es der Benutzer betrachten kann, vorzugsweise im Untertitelformat oder im Bildschirmformat. Dieses Gerät zur visuellen Wiedergabe gesprochener Sprache kann auch ein Teilsystem zur Sprachgenerierung für die Benutzer umfassen, die vielleicht auch eine Sprachbehinderung haben. Das erfindungsgemäße Gerät zur visuellen Wiedergabe gesprochener Sprache kann als Abwandlung der verschiedenen Bildzugriffsgeräte angesehen werden oder auch als Sonderfall bzw. spezielle Medienanwendung des Multimediasteuergeräts gemäß den hiermit zusammenhängenden Anmeldungen gelten. Wie im folgenden noch ausführlicher beschrieben ist, umfaßt die Vorrichtung zur visuellen Wiedergabe gesprochener Sprache viele derselben Bauelemente und Teilsysteme der Bildzugriffsgeräte und des Multimediasteuergeräts, so daß hier hinsichtlich der entsprechenden ausführlichen Beschreibungen und der technischen Angaben zu den bevorzugten Bauelementen auf die hiermit zusammenhängenden Anmeldungen verwiesen werden kann.

Fig. 1 ist ein Blockschaltbild mit der Darstellung eines erfindungsgemäßen Geräts **101** zur visuellen Wiedergabe gesprochener Sprache und eines erfindungsgemäßen Systems **100** zur visuellen Sprachwiedergabe. Entsprechend der Darstellung in **Fig. 1** umfaßt das System **100** zur visuellen Sprachwiedergabe das Gerät **101** zur visuellen Sprachwiedergabe, sowie mindestens ein Bildschirmgerät **225** und mindestens eine physikalische Schnittstelle **155**, beispielsweise in Form eines Telefons **150**, einer Tastatur **160**, einer Maus **170** oder eines Rechners **175**. Das Gerät **101** zur visuellen Sprachwiedergabe läßt sich über eine Netzschnittstelle **110** an einen ersten Kommunikationskanal (bzw. einem Netzkommunikationskanal) **103** zur Kommunikation mit einem Netz **104** ankoppeln. Der erste bzw. Netz-Kommunikationskanal **103** wird hier auch als Netzkommunikationskanal 03 bezeichnet, um ihn von anderen Kommunikationskanälen des Systems **100** zur visuellen Sprachwiedergabe zu unterscheiden, z. B. vom zweiten Kommunikationskanal **227**, der zur Kommunikation mit den verschiedenen Bildschirmgeräten **225** verwendet wird, oder vom dritten Kommunikationskanal **228**, der zur Kommunikation mit der Tastatur **160** oder den anderen physikalischen Schnittstellen **155** verwendet wird. Der erste Kommunikationskanal **103** kann fest verdrahtet sein, z. B. kann er aus einem oder mehreren verdrillten Drahtpaaren bestehen, oder es kann sich dabei um ein Kabel handeln, z. B. ein hybrides Glasfaser-Koaxkabel, auch um eine schnurlose Verbindung wie sie beispielsweise bei Mobiltelefonen oder für andere Hochfrequenz-Übertragungen verwendet wird, oder auch um jedes andere geeignete Kommunikationsmedium. Das Netz **104** kann, wie in der hiermit zusammenhängenden vierten und fünften Anmeldung beschrieben, beispielsweise ein sogenanntes PSTN-Netz (öffentliches Telefonnetz mit Wählbetrieb) für den normalen Telefonverkehr und analogen Datenverkehr sein (diese Einsatzgebiete werden hier als POTS bezeichnet; Plain Old Telephone Service, einfacher alter Telefondienst) oder ein sogenanntes ISDN-Netz für die digitale Übermittlung von Sprache und Daten, oder es kann sich dabei um eine Kombination aus derartigen schon bestehenden oder künftigen Telekommunikationsnetzen handeln. Bei solchen festverdrahteten Netzen wird der Netzkommunikationskanal **103** normalerweise über eine lokale digitale oder analoge (hier nicht dargestellte) Weiche an das Netz **104** angekoppelt. Außerdem kann, wie in der hiermit zusammenhängenden zweiten und dritten Anmeldung beschrieben, der Netzkommunikationskanal **103** bei Realisierung eines CACS-Kommunikationsprotokolls über eine Primärstation an das Netz **104** angekoppelt werden, die unter anderem mindestens eine Netzschnittstelle bietet, die mit anderen oder zusätzlichen Protokollen – z. B. den verschiedenen ISDN-Protokollen – arbeitet und außerdem für den Anschluß an eine Infrastruktur für Kabelfernsehdienste (CATV-Dienste) sorgt.

Das in **Fig. 1** dargestellte Gerät **101** zur visuellen Sprachwiedergabe weist verschiedene Merkmale oder Bauelemente auf, die in den hiermit zusammenhängenden Anmeldungen ausführlich beschrieben werden, unter anderem die Netzschnittstelle **110**, die Prozessorengruppe **130** (wobei verstanden wird, daß eine Prozessorengruppe auch nur einen Prozessor aufweisen kann) und die Benutzerschnittstelle **120**. In den zugehörigen Anmeldungen sind außerdem detaillierte Blockschaltbilder und technische Angaben zu den bevorzugten Bauelementen enthalten. Je nach der jeweiligen Realisierungsform des Systems **100** zur visuellen Sprachwiedergabe, z. B. in festverdrahteter, verkabelter oder schnurloser Form, ist die Netzschnittstelle **110** des Geräts **101** zur visuellen Sprachwiedergabe unterschiedlich aufgebaut. Bei Kabeltechnik ist beispielsweise die Netzschnittstelle **110** eine Kabelnetzschnittstelle mit einem

CATV-Sende-Empfänger und einer ASIC-Schaltung für Kommunikationszwecke (anwendungsspezifische integrierte Schaltung), die verschiedene Funktionen erfüllt wie zum Beispiel jeweils die Hochfrequenzmodulation (HF-Modulation) und -Demodulation und die Kodierung und Dekodierung nach dem CACS-Protokoll, wie in der zugehörigen zweiten und dritten Anmeldung beschrieben ist. Bei schnurlosen Einsatzgebieten, z. B. gemäß der Beschreibung in der zugehörigen vierten und fünften Anmeldung, umfaßt die Netzschnittstelle **110** eine Telefonschnittstelle (POTS-Schnittstelle für den traditionellen Telefonverkehr) und/oder eine ISDN-Schnittstelle, die jeweils verschiedene Funktionen hat, z. B. jeweils die analoge Telefontechnik (und ebenso analoge Modemfunktionen, z. B. nach den ITU-Protokollen (International Telecommunications Union) V.34 und V.34^2), neben der Realisierung verschiedener digitaler (ISDN-) Protokolle für Sprach- und Datenübermittlung (z. B. Protokolle zur digitalisierten Datenverbindung nach ITU Q.921 LAPD und für physikalische Layers (Interface-Protokolle) nach Q.910). Gemäß den zugehörigen Anmeldungen wird die Netzschnittstelle **110** zur Übertragung und zum Empfang analoger oder digitaler Bild-, Ton- und anderer Informationen und Daten (die ganz allgemein hier als Daten bezeichnet werden) in jedem gegebenen Format, mit jedem Protokoll oder nach jedem Modulationsschema eingesetzt, die mit dem Netz **104** kompatibel sind, wobei auch jede beliebige Form des Netzanschlusses oder der Schaltungen Verwendung findet. Wenn beispielsweise über den ersten Kommunikationskanal **103** der Anschluß an ein digitales Netz (z. B. ISDN-Netz) vorgesehen ist, übermittelt und empfängt die Netzschnittstelle **110** Daten in Form eines Tonsignals für den Telefonverkehr, oder als gemäß der ISDN-Protokollserie (z. B. Serie Q.x) kodierte und formatierte digitale Information. Bei Anschluß an ein herkömmliches bzw. PSTN-Netz über den ersten Netzkommunikationskanal **103** übermittelt und empfängt die Netzschnittstelle **110** beispielsweise auch Daten wie Tonsignale, z. B. ein normales analoges Tonsignal in POTS-Technik.

Aus **Fig.** 1 ist weiterhin ersichtlich, daß an die Netzschnittstelle **110**, an eine Benutzerschnittstelle **120** und an einen Hochfrequenz- bzw. HF-Modulator **270** eine Prozessorengruppe **130** angeschlossen ist. Die Netzschnittstelle **110**, die Benutzerschnittstelle **120** und der HF-Modulator **270** sind im wesentlichen identisch mit den Gruppen ausgelegt, wie sie in den zugehörigen Anmeldungen beschrieben und ausführlich dargestellt sind. Verschiedene Funktionen jeder dieser Systemkomponenten werden nachstehend außerdem noch ausführlicher erläutert. Bei dem in **Fig.** 1 dargestellten Ausführungsbeispiel weist die Vorrichtung **101** zur visuellen Sprachwiedergabe beispielsweise zunächst eine Netzschnittstelle **110** auf, die sich zum Empfangen eines ersten Tonsignals von einem Netz **104** an einen ersten Kommunikationskanal **103** ankoppeln läßt woraufhin sie aus diesem Signal ein Tonempfangssignal bildet; zum anderen weist sie einen Hochfrequenzmodulator **270** auf, der ein Bildausgangssignal (aus der Prozessorengruppe **130**) im Basisband in ein Hochfrequenz-Bildausgangssignal umwandelt und das Hochfrequenz-Bildausgangssignal zu einem zweiten Kommunikationskanal **227** zur Videodarstellung überträgt, z. B. über eines der Bildschirmgeräte **225**; und zum dritten umfaßt sie eine Prozessorengruppe **130**, die an die Netzschnittstelle **110** und den Hochfrequenzmodulator **270** angekoppelt ist und über einen Satz Programmbefehle in nachstehend noch erläuterter Weise so angesteuert wird, daß sie das empfangene Tonsignal in eine Textdarstellung der gesprochenen Sprache umsetzt und außerdem die Textdarstellung der Sprache in ein Bildausgangssignal im Basisband umwandelt (das dann vom HF-Modulator **270** noch moduliert und dann übertragen werden muß). Im folgenden wird außerdem noch erläutert, daß die Vorrichtung zur visuellen Sprachwiedergabe vorzugsweise die Benutzerschnittstelle **120** zur Eingabe von Steuersignalen umfaßt, die zur Ansteuerung verschiedener Betriebsarten – z. B. normaler Telefonbetrieb oder Modus mit visueller Sprachwiedergabe – verwendet werden.

Die Benutzerschnittstelle **120** dient zum Empfangen eines Steuersignals von aus einer Vielzahl von Steuersignalen, z. B. in Form einer Anforderung für ein Telefongespräch, einer Anforderung für die visuelle Darstellung der gesprochenen Sprache während eines Telefongesprächs, oder ein Anruf in einer Audio-/Video-Konferenz, einer Anforderung von Sprachgenerierung aus einem eingegebenen Text, und weitere Steuersignale wie zum Beispiel Meldesignale zur Ankündigung eines ankommenden Anrufs oder von Anrufen bei einer audiovisuellen Konferenz. Bei dem bevorzugten Ausführungsbeispiel ist die Benutzerschnittstelle **120** in Form einer Benutzer-Tonschnittstelle **255** ausgeführt, wie sie in **Fig.** 2 und 3 und ausführlich in den hiermit zusammenhängenden Anmeldungen dargestellt ist. Der HF-Modulator **270** setzt ein Bildausgangssignal in ein Hochfrequenz-Bildausgangssignal um, wie in den zugehörigen Anmeldungen beschrieben und dargestellt, und überträgt dieses zum zweiten Kommunikationskanal **227** und bringt es zur Anzeige auf den Bildschirmgeräten **225**. Bei dem bevorzugten Ausführungsbeispiel handelt es sich bei dem zweiten Kommunikationskanal **227** um ein Koaxkabel, wie es für Kabelfernsehen vorgesehen ist und im Raum beim Benutzer bzw. Teilnehmer an einer oder mehreren Stellen verlegt ist.

Die Prozessorengruppe **130** sorgt für die Umsetzung des empfangenen Tonsignals (aus der Netzschnittstelle **110**) in eine visuelle Darstellung der gesprochenen Sprache bzw. in deren Darstellung in Textform, die ihrerseits dann in die Form des Bildausgangssignals im Basisband umgesetzt wird (das vom HF-Modulator **270** noch moduliert und an die Bildschirmgeräte **225** übertragen werden muß. Die Prozessorengruppe **130** kann auch für die Sprachgenerierung aus einem eingegebenen Text sorgen (wobei die Sprachsignale dann über die Netzschnittstelle **110** an das Netz **104** übermittelt werden sollen). Wie in den zugehörigen Anmeldungen dargestellt und nachstehend noch ausführlicher erläutert wird, kann die Prozessorengruppe **130** aus einer einzigen integrierten Schaltung ("IC") bestehen oder eine Vielzahl integrierter Schaltungen bzw. anderer Bauelemente aufweisen, die miteinander verbunden bzw. zu Gruppen zusammengefaßt sind, z. B. Mikroprozessoren, digitale Signalprozessoren, ASIC-Schaltungen, zugehörige Speicher (z. B. RAM- und ROM-Speicher) und weitere ICs und Baugruppen. Infolgedessen ist der hier verwendete Begriff "Prozessorengruppe" als gleichbedeutend mit einem einzelnen Prozessor oder mit einer Anordnung von Prozessoren, Mikroprozessoren, Steuerungen oder irgendwelchen anderen Gruppierungen integrierter Schaltungen zu verstehen, welche die nachstehend noch näher beschriebenen Funktionen ausführen. Bei dem bevorzugten Ausführungsbeispiel ist zum Beispiel die Prozessorengruppe **130** gemäß Darstellung in **Fig.** 2 und 3 als Mikroprozessor-Teilsystem **260** ausgeführt (wie sie auch in den zugehörigen Anmeldungen dargestellt wird), neben einem Teilsystem zur visuellen Sprachwiedergabe (**300** bzw. **310**) und. kann außerdem ein Teilsystem zur Sprachgenerierung (**320**) umfassen.

Aus **Fig.** 1 ist weiterhin zu entnehmen, daß das Gerät **101** zur visuellen Sprachwiedergabe über die Benutzerschnittstelle **120** an mindestens eine physikalische Schnittstelle **155** gekoppelt ist, damit der Benutzer zur Eingabe eines oder mehrerer Steuersignale und auch für die Eingabe von Text zur Sprachgenerierung physikalischen Zugang zu der

Vorrichtung zur visuellen Sprachwiedergabe hat. Die physikalischen Schnittstellen **155** umfassen im typischen Fall mindestens ein Telefon **150**, eine Tastatur **160**, eine Computermaus **170** oder einen Rechner **175**. Die Telefone **150** können auch als Bildtelefon ausgeführt sein. Sind Telefone **150** in das System geschaltet, so erfolgt die physikalische Eingabe der Vielzahl von Steuersignalen über eine Telefontastatur in Form eines DTMF-Signals (Zweiton-Mehrfrequenzsignal) oder Impulswahlsignals, wobei für den normalen Eingang und Ausgang der Tonsignale eine Sprechmuschel und ein Hörerteil bei den verschiedenen Telefonen **150** (bzw. Bildtelefonen) vorgesehen sind. Zusätzlich zu den Telefonen **150**, oder auch anstelle derselben, können auch die Tastatur **160**, die Maus **170**, und/oder der Rechner **175** zur Eingabe der Vielzahl von Steuersignalen eingesetzt werden. Die Tastatur **160** bzw. der Rechner **175** dienen vorzugsweise für die Eingabe eines Textes für die Sprachgenerierung über den dritten Kommunikationskanal **228** (auch wenn andere Eingabeverfahren wie beispielsweise das DTMF-Wählverfahren ebenfalls herangezogen werden könnten). Der dritte Kommunikationskanal **228** wird hier als Kanal mit direkter Verbindung zwischen den physikalischen Schnittstellen **155** und der Prozessorengruppe **130** dargestellt, auch wenn andere Möglichkeiten der Verbindung zur Verfügung stehen; beispielsweise kann der dritte Kommunikationskanal **228** auch völlig entfallen (**Fig.** 2), wobei dann die Eingabe der Steuersignale über eine Verbindung (z. B. Leitung **294** in **Fig.** 2) mit der Benutzerschnittstelle **120** oder eine Benutzer-Tonschnittstelle **25** (statt mit der Prozessorengruppe **130**) erfolgt.

Fig. 1 zeigt weiterhin, daß der HF-Modulator **270** ein Bildausgangssignal im Basisband von der Prozessorengruppe **130** – z. B. in Form eines kombinierten NTSC/PAL-Videosignals – in ein Hochfrequenz-Bildausgangssignal umsetzt, z. B. ein amplitudenmoduliertes Restseitenband-HF-Signal, das über ein Bildschirmgerät **225** betrachtet werden kann, oder, wie **Fig.** 2 und 3 dies zeigen, beispielsweise über ein Fernsehgerät **240** des Benutzers, wenn dieser auf Kanal **3** oder **4** eingestellt wird. Der HF-Modulator **270** kann auf vielerlei Weise realisiert werden, unter anderem unter Verwendung eines Bildmodulators, z. B. Motorola MC1373, an den sich eine Verstärkungsstufe anschließt, die bei dem bevorzugten Ausführungsbeispiel dazu eingesetzt wird, Verluste aus einem Richtkoppler **290** (in **Fig.** 2 dargestellt) auszugleichen, die gegebenenfalls das HF-Bildausgangssignal in den zweiten Kommunikationskanal **227** einspeisen, z. B. in das Koaxkabelsystem in den Räumen des Benutzers.

Wie nachstehend noch ausführlicher erläutert wird, läßt sich die erfindungsgemäße Verfahrensweise in Form eines Satzes Programmbefehle zur anschließenden Ausführung in der Prozessorengruppe **130** und dem zugehörigen Speicher und anderen äquivalenten Bauelementen programmieren und abspeichern. Der Satz Programmbefehle kann auch in jeder Speichereinrichtung abgelegt werden, z. B. in Form eines Speicherbausteins in Form einer integrierten Schaltung, einer Diskette, einer CD-ROM oder in Form jedes anderen lesbaren oder abarbeitbaren Mediums. Bei dem bevorzugten Ausführungsbeispiel wird die Prozessorengruppe **130** in Verbindung mit einem abgespeicherten Satz Programmanweisungen und im Ansprechen auf alle vom Benutzer eingegebenen oder aus dem Netz **104** empfangenen Steuersignale für viele verschiedene Funktionen eingesetzt werden. Infolgedessen weist das bevorzugte Ausführungsbeispiel der Prozessorengruppe **130** eine Vielzahl von Betriebsarten auf, z. B. Betriebsarten zur visuellen Sprachwiedergabe, für normalen Telefonbetrieb (POTS-Betrieb), für die Übermittlung synthetisierter Sprache und auch bei Audio- und Video-Konferenzen (bei einem bevorzugten Ausführungsbeispiel).

Das Blockschaltbild in **Fig.** 2 zeigt ein erstes bevorzugtes Ausführungsbeispiel einer Vorrichtung **201** zur visuellen Sprachwiedergabe sowie ein erstes bevorzugtes Ausführungsbeispiel eines Systems **200** zur visuellen Sprachwiedergabe, beide gemäß der vorliegenden Erfindung. Das System **200** zur visuellen Sprachwiedergabe **200** umfaßt ein Gerät **201** zur visuellen Sprachwiedergabe **201**, mindestens ein Telefon **150** (in Form der physikalischen Schnittstellen **155**) und mindestens ein Fernsehgerät **240** (als eine Art des Bildschirmgeräts **225**), die über den zweiten Kommunikationskanal **227** mit der Vorrichtung **201** zur visuellen Sprachwiedergabe gekoppelt sind. Über die vorstehend angesprochene Netzschnittstelle **110** läßt sich die Vorrichtung **201** zur visuellen Sprachwiedergabe auch an ein (hier nicht dargestelltes) Netz **104** ankoppeln. Die Vorrichtung **201** zur visuellen Sprachwiedergabe umfaßt außerdem einen HF-Modulator **270**, der mit einem Richtkoppler **290** gekoppelt ist, der in vorstehend erläuterter Weise das HF-Bildausgangssignal vom HF-Modulator **270** in den zweiten Kommunikationskanal **227** überträgt, beispielsweise in Form einer in den Räumen des Benutzers verlegten Koaxkabelanlage.

Wie in den hierzu gehörigen Anmeldungen im einzelnen erläutert ist, wird die Benutzertonschnittstelle **255** in der Weise ausgelegt, daß sie den Übergang zu üblichen Haushaltstelefonapparaten bildet, worunter auch schnurlose Apparate und Freisprechgeräte wie die Telefone **150** fallen. Die Benutzertonschnittstelle **255** soll sowohl für bisher übliche Gespräche in POTS-Technik als auch für Bildtelefonie geeignet sein und in Verbindung mit der Netzschnittstelle **110** auch analoge Modemfunktionen unterstützen. Darüber hinaus sorgt die Benutzertonschnittstelle in Verbindung mit einer der physikalischen Schnittstellen **155** – beispielsweise dem Telefon **150** (bzw. der Tastatur **160**, der Maus **170** oder dem Rechner **175**, die in **Fig.** 1 dargestellt sind) – für die Eingabe der verschiedenen Steuersignale, wie sie beispielsweise zur Anwahl einer Anwendung mit visueller Sprachwiedergabe oder zur Telefonanwahl oder Bildtelefonanwahl verwendet werden. Bei dem bevorzugten Ausführungsbeispiel wird jedes der Telefone **150** zur Eingabe der verschiedenen Steuersignale verwendet, und Anrufe in normaler POTS-Technik werden in "transparenter" Form verarbeitet, was bedeutet, daß ausgehende und ankommende Telefonanrufe so ablaufen, als ob die Funktionen zur visuellen Sprachwiedergabe, für Videokonferenz oder andere Multimediafunktionen nicht vorhanden wären. Darüber hinaus werden bei dem bevorzugten Ausführungsbeispiel die Funktionen mit visueller Sprachwiedergabe, mit Bildtelefonanrufen und Multimediafunktionen als Ausnahmefälle bearbeitet, wobei der Benutzer eine jeweilige spezielle bzw. vorgegebene Wählfolge eingeben muß, um die visuelle Sprachwiedergabe, einen Bildtelefonanruf oder eine andere Medienfunktion anzusteuern. Die bei dem bevorzugten Ausführungsbeispiel verwendeten verschiedenen Telefone **150** können in jeder Art normaler Telefone ausgeführt sein, einschließlich schnurloser (tragbarer) Telefone, der üblichen Telefone mit Schnuranschluß, DTMF- oder Impulswahltelefone, Bildtelefone oder Freisprechtelefone.

Wie auch in den zugehörigen Anmeldungen beschrieben, weist die Benutzertonschnittstelle **255** vorzugsweise eine SLIC-Schaltung auf (Subscriber Loop Interface Circuit; Teilnehmer-Schleifenschnittstellenschaltung), die sogenannte "BORSHT"-Funktionen für Telefondienste innerhalb der Räume des Benutzers bietet, sowie eine Ringbildungsschaltung; einen Tonkodierer/-dekodierer für den Audioanteil eines Bildtelefongesprächs oder normalen Telefongesprächs, wobei dieses Teil für die Analog-Digital-Umsetzungen zur Digitalisierung eines Sprach- bzw. Tonein-

gangssignals aus dem Sprechmuschelteil eines oder mehrerer der Telefone **150** und die Digital-Analog-Umsetzung für die Stimmenwiedergabe aus den Datenstrom bzw. Signal eines digitalisierten Sprachausgangssignals sorgt (um ein Tonausgangssignal daraus zu bilden, das dem Sprechteil der Telefone **150** zugeleitet wird); und schließlich einen programmierbaren digitalen Signalprozessor (DSP) mit zugehörigem Speicher (der als DSP zur Stimmverarbeitung in den zugehörigen Anmeldungen bezeichnet wird, im Unterschied zu einem anderen DSP-Element, das als DSP-Teil zur Bildverarbeitung bezeichnet wird). Das DSP-Element in der Benutzertonschnittstelle **255** enthält einen Programmspeicher und einen Datenspeicher zur Ausführung von Funktionen zur Signalverarbeitung, beispielsweise die Erfassung von DTMF/Wählimpulswahl und Impulserzeugung, für analoge Modemfunktionen, zur Bildung von Rufablauftönen (Wählton, Belegttonzeichen), für die PCM-Linear-Umsetzung und die Linear-PCM-Umsetzung (Impulskodemodulation) und die Abspielung der Sprechaufforderung. Der dem DSP-zugeordnete Speicher weist bei dem bevorzugten Ausführungsbeispiel einen Festspeicher (als Sprach-ROM-Speicher bezeichnet) hoher Dichte mit PCM-kodierten (bzw. komprimierten) Sprachsegmenten auf, die zur Interaktion mit dem Benutzer verwendet werden, z. B. zur Aufforderung des Benutzers zur Eingabe des DTMF- oder Impulswahlverfahrens über die Tastatur, wenn die Verbindung über die Bildtelefoniefunktion oder eine in einem anderen Multimediamodus hergestellt werden soll. Daneben kann bei Bedarf ein Sprach-RAM-Speicher für Speicherfunktionen zur Sprachspeicherung durch den Benutzer verwendet werden, sowie ein elektrisch veränderbarer programmierbarer leistungsloser (schnell löschbarer) Speicher zur Abspeicherung von Programmen (und Programmerweiterungen) oder Algorithmen.

Die Prozessorengruppe **130** (gemäß **Fig.** 1) ist bei dem Gerät **201** zur visuellen Sprachwiedergabe in Form eines Mikroprozessor-Teilsystems **260** und eines Teilsystems (bzw. Prozessors) **305** zur visuellen Sprachwiedergabe ausgeführt, der in **Fig.** 2 dargestellt ist. Wie ausführlich in den hierzu gehörigen Anmeldungen dargestellt, besteht das Mikroprozessor-Teilsystem **260** aus einem Mikroprozessor oder einer anderen Verarbeitungseinheit, beispielsweise in Form des Motorola-Bauteils MC68LC302, und einem Speicher, der einen Direktzugriffsspeicher (RAM) und einen Festwertspeicher (ROM) und bei dem bevorzugten Ausführungsbeispiel auch einen sogenannten programmierbaren Flash-Speicher (z. B. Flash-EPROM bzw. E^2PROM) umfaßt, wobei die Kommunikationsverbindung über die Busleitung **261** mit der Netzschnittstelle **110**, der Benutzertonschnittstelle **255** und über die Busleitung **263** mit dem Teilsystem **305** zur visuellen Sprachwiedergabe hergestellt wird. Für den Festwertspeicher wird ebenfalls ein programmierbarer Flash-Speicher verwendet, so daß der Speicherinhalt aus dem Netz **104** heruntergeladen werden kann. Infolgedessen können verschiedene Versionen der Betriebssoftware (Programmbefehle) wie z. B. Programmverbesserungen realisiert werden, ohne daß an dem Gerät **201** zur visuellen Sprachwiedergabe Veränderungen vorgenommen werden müssen und ohne daß der Benutzer eingreifen muß. Das Mikroprozessor-Teilsystem **260** sorgt für die Steuerung und Konfigurierung des Prozessors **305** zur visuellen Sprachwiedergabe, für die Verarbeitung normaler Telefongespräche, von Telefongesprächen in Digitaltechnik und wird außerdem zur Implementierung eines ISDN-Stapels oder eines anderen Protokollstapels eingesetzt, wenn dies für analoge oder digitale Bildtelefonieverbindungen erforderlich ist, z. B. bei Meldungsübermittlung mit ITU Q.931-Protokoll.

Das Teilsystem **305** zur visuellen Sprachwiedergabe, das auch als Prozessor zur visuellen Sprachwiedergabe bezeichnet wird, kann außerdem aus einem Mikroprozessor bzw. einer anderen Verarbeitungseinheit wie beispielsweise dem Motorola-Baustein MC68LC302 und aus einem Speicher bestehen, der bei dem bevorzugten Ausführungsbeispiel einen RAM- und einen ROM-Speicher und außerdem einen programmierbaren Flash-Speicher (z. B. Flash-EPROM bzw. E^2PROM) umfaßt. Wie **Fig.** 2 zeigt, gehören zu dem Teilsystem **305** zur visuellen Sprachwiedergabe auch zwei Funktionsblöcke, und zwar ein Teilsystem (bzw. Prozessor) **307** zur Spracherkennung und ein Teilsystem (bzw. Prozessor) **309** zur Bildschirmdarstellung.

Je nach Art der Netzschnittstelle **110** und dem zugehörigen bzw. entsprechenden Netz **104** können die aus dem Netz **104** ankommenden Sprachsignale verschiedene Formate aufweisen. Beispielsweise werden bei Anschluß an das PSTN-Netz die ankommenden Sprachsignale in Form analoger Signale von der Netzschnittstelle **110** empfangen und vorzugsweise in ein digitales Format umgewandelt, z. B. in ein impulskodemoduliertes (PCM) digitales Sprachsignal. Bei Anschluß an ein Kabelnetz werden die ankommenden Sprachsignale von der Netzschnittstelle **110** als CACS-Signale oder als Signale nach einem anderen Empfangsprotokoll empfangen, die dann zur Bildung eines digital kodierten Sprachsignals, z. B. eines PCM-kodierten Sprachsignals, demoduliert werden können. Sind die Sprachsignale Teil einer Audio-/Video-Konferenz, so trennt das Mikroprozessor-Teilsystem **260** das digitale Sprachsignal vom Bildsignalanteil zur separaten Verarbeitung (wie nachstehend anhand von **Fig.** 3 erläutert wird). Das digitale Sprachsignal wird dann zum Teilsystem **307** zur Spracherkennung übermittelt. Bei dem bevorzugten Ausführungsbeispiel ist das Teilsystem **307** zur Spracherkennung mit einer Spracherkennungs-Software programmiert, die eine Eigenentwicklung oder auch eine im Handel erhältliche Software sein kann, z. B. das Softwaresystem zur Spracherkennung von IBM oder Lexicus (einer Tochtergesellschaft von Motorola, Inc.). Bei dem bevorzugten Ausführungsbeispiel kann das Spracherkennungs-Teilsystem **307** im Laufe der Zeit trainiert werden, um so die Präzision bei der Spracherkennung bei häufigen Anrufern zu erhöhen. Das Teilsystem **307** zur Spracherkennung generiert aus dem digitalen Sprachsignal eine Textdarstellung der gesprochenen Sprache, die unterschiedlich formatiert sein kann, d. h. in Form eines Textes im ASCII-Format oder eines Textes in einer anderen entsprechend kodierten bzw. formatierten Form. Die Textdarstellung der gesprochenen Sprache wird dann zum Teilsystem **309** zur Bildschirmdarstellung übertragen, das ebenfalls mit einer handelsüblichen oder speziell entwickelten Software programmiert ist. Das Teilsystem **309** zur Bildschirmdarstellung kann auch unter Verwendung einer separaten integrierten Schaltung, z. B. OSD PCA855D von Philips, realisiert werden. Das Teilsystem **309** zur Bildschirmdarstellung setzt die Textdarstellung der gesprochenen Sprache in ein Format zur Bildschirmdarstellung um, das dann als Bildausgangssignal im Basisband an den HF-Modulator **270** ausgegeben wird. Es können auch andere Bildformate herangezogen werden, z. B. das nachstehend anhand von **Fig.** 3 erläuterte Untertitelformat. Der HF-Modulator setzt das Bildausgangssignal im Basisband in ein Hochfrequenz-Bildausgangssignal um, das dann beispielsweise auf Kanal **3** oder **4** über den zweiten Kommunikationskanal **227** zur Anzeige auf den verschiedenen Fernsehgeräten **240** übertragen wird. Infolgedessen setzt das Gerät **201** zur visuellen Sprachwiedergabe das empfangene Tonsignal – z. B. in Form eines aus einem Netz kommenden Sprachsignals – in ein Hochfrequenz-Bildausgangssignal um, das dann zu einem oder mehreren Bildschirmgeräten (z. B. zum Fernsehgerät **240**)

zur visuellen Anzeige des gesprochenen Textes übertragen wird.

Wie vorstehend und auch in den hierzu gehörigen Anmeldungen dargestellt, leitet der Benutzer bei dem bevorzugten Ausführungsbeispiel einen Multimedia-Modus– z. B. die visuelle Darstellung des Sprachmodus oder eines Videokonferenzmodus – dadurch ein, daß er im Unterschied zum normalen bzw. gewöhnlichen Telefonbetrieb eine spezielle vorgegebene Folge eintippt, die von dem DSP-Element in der Benutzertonschnittstelle **255** als Folge für einen Multimedia-Modus erkannt wird. Alternativ kann eine Vielzahl von Signalfolgen für einen Multimediamodus verwendet werden, wobei jede vorgegebene Folge speziell für einen gewählten Multimediamodus gilt, z. B. Videomodus oder Betriebsart mit visueller Sprachwiedergabe. Diese Methodik wird auch im folgenden anhand des Ablaufdiagramms in **Fig.** 4 noch erläutert. Für einen Multimedia-Modus sind bei dem bevorzugten Ausführungsbeispiel die ersten beiden Ziffern der spezifischen vorgegebenen Folge beispielsweise nur für diese Anwendung reserviert und werden bei einem üblichen Anruf in POTS-Technik nicht verwendet, z. B. "**"; infolgedessen können sie speziell dem DSP-Element mitteilen, daß statt eines normalen Telefonbetriebsmodus nun in einen Multimedia-Modus umgeschaltet werden soll. Alternativ könnte der Benutzer auch andere spezifische vorgegebene Folgen als Kennung für einen Multimedia-Modus einprogrammieren. Die verschiedenen Medien-Betriebsarten könnten lokal über eine der physikalischen Schnittstellen **155** oder aus der Entfernung über einen Anschluß über das Netz **104** und die Netzschnittstelle eingegeben werden. Unmittelbar nach Dekodierung der beiden speziellen Ziffern oder einer anderen spezifischen vorgegebenen Folge als Hinweis auf einen Multimedia-Modus leitet das Gerät **201** zur visuellen Sprachwiedergabe den Ablauf zur Ansteuerung der visuellen Sprachwiedergabe (bzw. der Multimedia-Anwendung) ein, beispielsweise dadurch, daß das DSP-Element eine Abfolge zur Aufforderung in gesprochener Sprache oder per Videodarstellung generiert, abspielt oder anzeigt, z. B. "Bitte wählen Sie eine Anrufoption oder drücken Sie Taste "#" zur Hilfestellung", die im ROM-Teil des Speichers in der Benutzertonschnittstelle **255** abgespeichert ist. Was das DSP-Element unternimmt erfolgt dann im Ansprechen auf die eingegebene Folge oder auf den Tastendruck des Benutzers nach der ersten Aufforderung und hängt davon ab. Wird beispielsweise die Taste "#" gedrückt, kann der Benutzer ein Befehlsmenü sehen oder hören, zum Beispiel in dieser Form:

- "zur Eingabe der visuellen Darstellung gesprochener Sprache – 1 drücken"
- "zur Eingabe des Videokonferenz-Modus – 2 drükken"
- "zur Eingabe der Automatisierung im Haus – 3 drükken"
- "zur Eingabe gesprochener Nachrichten – 4 drükken"
- "zum nochmaligen Abspielen dieses Menüs – # drücken".

Nach Auswahl des besonderen oder speziellen Medien-Modus durch den Benutzer, z. B. Bilddarstellung gesprochener Sprache, generieren das Gerät **201** bzw. das System **200** zur visuellen Sprachwiedergabe ein Untermenü mit Befehlen oder bringen dies auf den Bildschirm. Hat beispielsweise der Benutzer eine Betriebsart mit visueller Anzeige gesprochener Sprache gewählt, kann er ein Untermenü mit Befehlen sehen oder hören, wie zum Beispiel das folgende:

- "für Anruf über Rufnummernverzeichnis – * drükken"
- "für Aktualisierung des Rufnummernverzeichnisses – 2 drücken"
- "für manuellen Bildtelefonanruf – 3 drücken"
- "für Zuschaltung der Sprachgenerierung – 4 drükken"
- "zum nochmaligen Abspielen dieses Menüs – # drücken".

Einer der Vorteile des Rufnummernverzeichnisses des Benutzers besteht bei dem bevorzugten Ausführungsbeispiel darin, daß durch die Vorauswahl des anzurufenden Teilnehmers dem Teilsystem **307** für die Spracherkennung mitgeteilt werden kann, daß ein Teilnehmer angerufen werden soll, den es bereits "gelernt" hat, also ein Teilnehmer, mit dem das Teilsystem **307** zur Spracherkennung bereits ein gewisses Training absolviert hat. Infolgedessen kann das Teilsystem **307** zur Spracherkennung im wesentlichen noch feiner abgestimmt werden, um die Sprechweise einer bestimmten Person zu erkennen, wodurch die Präzision bei der visuellen Darstellung der hörbaren Sprache noch verbessert wird. Außerdem kann der Benutzer auch durch Eingabe dieser unterschiedlichen Steuersignale bei ankommenden Gesprächen dem Teilsystem **307** zur Spracherkennung einen Hinweis darauf übermitteln, daß ein bestimmter Teilnehmer angerufen hat, und zwar wiederum zur Aktivierung dieser Feinabstimmung des Teilsystems **307** zur Spracherkennung auf ein zuvor im Zusammenhang mit dem anrufenden anderen Teilnehmer erlerntes Muster.

Damit wird bei dem bevorzugten Ausführungsbeispiel eine automatisierte benutzerfreundliche Abfolge von Aufforderungen verwendet, um den Benutzer durch den Ablauf bzw. die Sequenz zur visuellen Sprachwiedergabe über eine einzige (bzw. integrierte) physikalische Schnittstelle, z. B. ein Telefon **150**, zu führen, statt über mehrere und unterschiedliche (und außerdem häufig verwirrende) Schnittstellen. Zu weiteren noch besser entwickelten Systemen zur Interaktion mit dem Benutzer können auch die Benutzung des Fernsehgeräts **240** oder eines anderen Bildschirmgeräts zur visuellen Bildschirmdarstellung eines Menüs mit Optionen gehören, wobei die Steuersignale vom Benutzer entsprechend eingegeben werden, z. B. als Anrufsteuerinformation oder als Informationen für einen vorzunehmenden Anruf, was auf unterschiedliche Weise geschehen kann, z. B. über die Tastatur auf den Telefonen **150**, über eine Verbindung zur Infrarot-Fernsteuerung mit dem Gerät **201** zur visuellen Sprachwiedergabe, oder mittels des zweiten Kommunikationskanals **227** (in **Fig.** 3 dargestellt) über einen Bildeingabepfad.

Das Blockschaltbild in **Fig.** 3 zeigt ein zweites bevorzugtes Ausführungsbeispiel des erfindungsgemäßen Geräts **301** zur visuellen Sprachwiedergabe und des erfindungsgemäßen Systems **300** zur visuellen Sprachwiedergabe (und Sprachgenerierung). Dabei umfaßt das System **300** zur visuellen Sprachwiedergabe (und Sprachgenerierung) ein Gerät **301** zur visuellen Sprachwiedergabe und Spracherkennung, mindestens ein Telefon **150** und eine Tastatur **160** (als physikalische Schnittstellen **155**), mindestens ein Fernsehgerät **240** (als eine Art Bildschirmgerät **225**), das über den zweiten Kommunikationskanal **227** mit dem Gerät **301** zur visuellen Sprachwiedergabe und Sprachgenerierung gekoppelt ist, eine Videokamera **230** und eine Kameraschnittstelle **235**. Die Videokamera **230** und die Kameraschnittstelle **235** werden in den hiermit zusammenhängenden Anmeldungen im einzelnen beschrieben und hier zum Zwecke der umfassenden Möglichkeit zur Videokonferenz herangezogen; dies geschieht in der Form, daß ein Video- bzw. Bildsignal aus der

Videokamera **230** und der Kameraschnittstelle **235** in den Räumen des Benutzers (durch den Demodulator **275**) demoduliert und (im Teilsystem **265** zur Audio-/Video-Kompression und -Dekompression) zur Übertragung durch das Gerät **301** zur visuellen Sprachwiedergabe und Sprachgenerierung über den ersten Kommunikationskanal **103** zu einem (hier nicht dargestellten) Netz **104** verarbeitet werden kann.

Aus **Fig. 3** ist des weiteren zu entnehmen, daß das Gerät **301** zur visuellen Sprachwiedergabe und Sprachgenerierung viele derselben Bauelemente und Baugruppen umfaßt, die vorstehend unter Bezugnahme auf **Fig. 2** erläutert wurde, z. B. eine Netzschnittstelle **110**, eine Benutzertonschnittstelle **255**, ein Mikroprozessor-Teilsystem **260**, einen HF-Modulator **270** und einen Richtkoppler **290**. Die Vorrichtung zur visuellen Sprachwiedergabe und Sprachgenerierung umfaßt einen zweiten Typus eines Teilsystems (Prozessors) zur visuellen Sprachwiedergabe, nämlich das Teilsystem (bzw. den Prozessor) **310** zur visuellen Sprachwiedergabe, der dazu eingesetzt wird, für die visuelle Darstellung gesprochener Sprache ein Untertitelformat zu bilden; des weiteren umfaßt das System auch ein Teilsystem (einen Prozessor) **320** zur Sprachgenerierung, welches eingegebenen Text in hörbare Sprachsignale zur Übertragung in das Netz **104** umsetzt. Die Vorrichtung zur visuellen Sprachwiedergabe und Sprachgenerierung ist außerdem mit mindestens einem Telefon **150** zur Eingabe von Steuersignalen und einer Tastatur **160** zur Texteingabe (für die anschließende Sprachgenerierung) gekoppelt. Das Gerät zur visuellen Sprachwiedergabe und Sprachgenerierung wird ebenfalls in der Weise gesteuert, wie sie vorstehend anhand des Geräts **201** zur visuellen Sprachwiedergabe erläutert wurde, und zwar durch Eingabe von Steuersignalen (vorzugsweise über ein Telefon **150**).

Wie in den zugehörigen Anmeldungen ausführlich dargestellt, führt das Teilsystem **265** zur Audio-/Video-Kompression und -Dekompression die Kompression und Dekompression von Ton- und Bildsignalen vor, vorzugsweise unter Verwendung von Protokollen aus der Serie ITU H.32x; dieses Teilsystem wird in erster Linie für Videokonferenzschaltungen eingesetzt. Für die visuelle Darstellung gesprochener Sprache aus dem Audioteil eines Videokonferenzanrufs (der über ein Netz 04 übertragen wird) dekomprimiert das Teilsystem **265** zur Audio-/Video-Kompression und -Dekompression das Tonsignal und trennt es vom Bildanteil des Videokonferenzanrufs ab. Dabei wird auch der Bildanteil des Videokonferenzanrufs dekomprimiert und in ein Bildausgangssignal im Basisband umgewandelt (was in den hierzu gehörenden Anmeldungen im einzelnen beschrieben wird). Das Tonsignal wird dann vom Teilsystem **307** zur Spracherkennung verarbeitet, um eine Darstellung der gesprochenen Sprache in Form eines geschriebenen Textes zu bilden, wie vorstehend anhand von **Fig. 2** erläutert wurde. Die Textdarstellung der gesprochenen Sprache wird dann vom Untertitel-Kodierer **311** verarbeitet, indem die Textdarstellung in ein Untertitelformat umgesetzt wird, was beispielsweise in der vertikalen Austastlücke geschehen kann. Der Untertitel-Kodierer **311** kann unter Verwendung eines handelsüblichen oder speziell hierfür entwickelten Untertitel-Kodierers bzw. Prozessors realisiert werden. Das Untertitel-Bildsignal im Basisband wird dann in einer Mischerstufe **313** mit dem Bildausgangssignal im Basisband (aus dem Bildteil des Videokonferenzanrufs) gemischt. Das gemischte Bildsignal, das nun die reine Bildinformation und die Untertitelinformation enthält, wird anschließend im HF-Modulator **270** zur Darstellung auf einem der Fernsehgeräte **240** moduliert und übertragen. Bei diesem Ausführungsbeispiel mit dem System **301** umfaßt ein Fernsehgerät **240** vorzugsweise einen Untertiteldekoder zur Dekodierung und Darstellung des Untertitelsignals.

Die zur Darstellung auf den verschiedenen Fernsehern oder anderen Bildschirmgeräten übertragenen Informationen zur visuellen Sprachwiedergabe können auch noch weitere Informationen enthalten. Beispielsweise läßt sich auch eine Lautstärkeinformation einbeziehen und darstellen, auch unter Verwendung einer Darstellung mit Sinuswellen zum Beispiel, wobei eine Amplitude mit der Lautstärke korreliert oder diese darstellt, oder unter Verwendung eines Fettdruck- oder Unterstreichungsformats, das ebenfalls mit der Lautstärke oder anderen Hervorhebungen in der gesprochenen Sprache korreliert.

Die Vorrichtung **301** zur visuellen Sprachwiedergabe und Sprachgenerierung umfaßt außerdem ein Teilsystem (einen Prozessor) **320** zur Sprachgenerierung, der mit einer Tastatur **160** zur Texteingabe für die anschließende Umsetzung in gesprochene Sprache und Übermittlung an ein Netz **104** gekoppelt ist. Bei dem bevorzugten Ausführungsbeispiel ist das Teilsystem **320** zur Sprachgenerierung, das auch als Sprachgenerator-Prozessor bezeichnet wird, mit einer Software zur Sprachgenerierung programmiert, die eine Sonderentwicklung für diesen Zweck oder eine handelsübliche Software sein kann oder unter Verwendung von handelsüblichen integrierten oder anderen Schaltungselementen realisierbar ist. Wie vorstehend im Hinblick auf ein ankommendes Signal in gesprochener Sprache erläutert wurde, kann das in das Netz **104** zu übertragende Ton- bzw. Sprachsignal je nach Art des Netzanschlusses unterschiedlich gebildet sein, wobei es sich zum Beispiel um ein analoges Tonsignal zu Übermittlung an ein PSTN-Netz, ein digitales Sprachsignal zur Übertragung in ein ISDN-Netz oder um ein Sprachsignal nach CACS-Protokoll zur Übertragung an eine Primärstation und anschließende Netzkommunikation handeln kann. Vorzugsweise wird zur Generierung von gesprochener Sprache Text über die Tastatur **160** in ein Teilsystem **321** zum Festhalten von Text beispielsweise in ASCII-Kodierung oder in anders kodierter oder auch binärer Form eingegeben und dann wird der Text aus diesem Format in Sprachformat umgesetzt (in Wörtern und Satzteilen), was in dem Teilsystem **322** zur Umsetzung von Text in Sprache geschieht. Das Sprachformatsignal wird dann in dem Sprachsynthesizer **323** in synthetisierte Sprache umgewandelt und kann danach in jedem geeigneten analogen, digitalen oder kodierten Format in ein Netz **104** übertragen werden.

Fig. 4 zeigt ein Ablaufdiagramm zur Veranschaulichung eines erfindungsgemäßen Verfahrens zur visuellen Sprachwiedergabe und zur Sprachgenerierung. **Fig. 4** zeigt dabei auch die verschiedenen Aufgaben bzw. Betriebsarten eines Telefons – z. B. des Telefons **150** – bei dem erfindungsgemäßen System auf, unter anderem für den normalen Telefonbetrieb (in POTS-Technik) zur für Multimedia-Steuerungszwecke, wozu auch Steuersignale zur Anwahl der Betriebsarten zur visuellen Sprachwiedergabe und zur Videokonferenzschaltung gehören. Gemäß **Fig. 4** beginnt das Verfahren mit dem Startschritt **400** und im Schritt **405** wird eine Bedienungsanforderung erfaßt, zum Beispiel Abheben oder Empfangen eines Meldesignals für einen ankommenden Anruf. Als nächstes erfolgt im Schritt **410** ein Hinweis bzw. eine Meldung an den Benutzer, z. B. mit visuell erkennbarem oder hörbarem Wählton, ein Läutesignal für einen ankommenden Anruf oder ein sichtbares Signal zur Meldung eines ankommenden Anrufs, und es werden Meldeinformationen zusammengefaßt z. B. DTMF-Ziffern für eine Telefonnummer oder "**". Wurde in Schritt **415** der Betriebsmodus zur visuellen Sprachdarstellung gewählt, z. B. durch Eingabe von "**" oder wird eine ankommende Meldung aus dem Netz **104** empfangen, verzweigt das Verfahren zum Schritt **435**. Wurde im Schritt **415** die Betriebsart zur visuel-

len Sprachdarstellung nicht angefordert, so läuft das Verfahren mit der Anforderung bzw. Anwahl eines normalen Telefongesprächs weiter, z. B. mit Generierung von DTMF-Tönen und Verbindung eines Audioschaltwegs zwischen dem Telefon des Benutzers und dem Netz **104** – Schritt **420** – woraufhin in den transparenten Telefonmodus geschaltet wird und Audiodaten (im typischen Fall PCM-Daten) im Schritt **425** zum Netz **104** übermittelt werden. Die Audiodaten wurden zuvor von der Benutzertonschnittstelle **255** PCM-kodiert und von der Netzschnittstelle **110** in ein entsprechendes digitales oder analoges Format (z. B. ISDN, POTS, etc.) zur Weiterleitung in das Netz **104** umgewandelt. Nach Beendigung des Telefongesprächs im Schritt **430** kann das Verfahren mit dem Rückkehrschritt **500** beendet sein.

Aus **Fig. 4** ist des weiteren ersichtlich, daß bei Anforderung der Betriebsart zur visuellen Sprachwiedergabe im Schritt **415** das Verfahren zum Schritt **435** verzweigt und nun feststellt, ob auch Sprachgenerierung angefordert wird. Wurde im Schritt **435** auch die Sprachgenerierung verlangt so verzweigt das Verfahren auch weiter zum Schritt **475** zur Sprachgenerierung gleichzeitig mit visueller Sprachdarstellung. Wurde im Schritt **415** unabhängig von der Anforderung von Sprachgenerierung im Schritt **435** nur die visuelle Sprachwiedergabe angefordert, so schaltet das Verfahren zum Schritt **440** weiter und initialisiert das System zur visuellen Sprachwiedergabe, zum Beispiel durch Abspielen einer einleitenden gesprochenen oder visuell dargestellten Aufforderung, wie vorstehend bereits erläutert wurde. Als nächstes wird im Schritt **445** ein Tonsignal empfangen, und das empfangene Tonsignal wird nun im Schritt **450** in eine Darstellung der gesprochenen Sprache in Textform umgewandelt. Die Textdarstellung der gesprochenen Sprache wird anschließend im Schritt **455** in ein Bildausgangssignal im Basisband umgewandelt und so moduliert, daß im Schritt **460** ein Hochfrequenz-Bildausgangssignal gebildet wird. Das Hochfrequenz-Bildausgangssignal wird anschließend im Schritt **465** zu einem Bildschirmgerät übertragen. Nach Beendigung des Schrilles der visuellen Sprachdarstellung im Schritt **470** kann das Verfahren zur visuellen Sprachwiedergabe mit dem Rückkehrschritt **500** beendet werden.

Wurde im Schritt **435** auch gleichzeitig mit dem Arbeitsgang zur visuellen Sprachwiedergabe in den vorstehend erläuterten Schritten **440** bis **470** die Sprachgenerierung angefordert, so verzweigt das Verfahren zum Schritt **475**, um das Teilsystem zur Sprachgenerierung zu initialisieren, was ebenfalls über die vorstehend dargestellten sichtbaren oder hörbaren Aufforderungen geschieht. Als nächstes wird im Schritt **480** eingegebener Text empfangen und im Schritt **485** wird der empfangene Eingabetext in ein Sprachsignal umgesetzt, das ein analog oder ein digital kodiertes Sprachsignal sein kann. Im Schritt **490** wird dann das Sprachsignal beispielsweise zu einem Telekommunikationsnetz übertragen; wenn dann der Arbeitsgang zur Sprachgenerierung im Schritt **495** beendet ist, kann das Verfahren mit dem Rückkehrschritt **500** beendet sein.

Zahlreiche Vorteile der verschiedenen erfindungsgemäßen Vorrichtungen, Verfahrensweisen und Systeme liegen klar auf der Hand. Zunächst sorgen die verschiedenen Geräte, Verfahren und Vorrichtungen gemäß der vorliegenden Erfindung für die visuelle Darstellung bzw. Wiedergabe von gesprochener Sprache, ohne daß lokal und am entfernt liegenden Ort bei einer Kommunikationsverbindung speziell nur für diesen Zweck vorgesehene Geräte und Systeme vorausgesetzt werden. Dabei kann jedes Telefon am entfernten bzw. weit abliegenden anderen Ende eingesetzt werden, wobei die übermittelten Informationen in gesprochener Sprache auf jedem angeschlossenen Fernsehgerät oder einem anderen Bildschirmgerät überall in den Räumen lokal beim Benutzer angezeigt werden können. Außerdem ist bei den verschiedenen Ausführungsbeispielen der vorliegenden Erfindung kein größerer Aufwand an manueller Betätigung für den Betrieb erforderlich. Beispielsweise ist es im Gegensatz zu Geräten nach dem Stand der Technik nicht erforderlich, den visuell darzustellenden Text über eine Tastatur einzugeben. Außerdem entfällt die Notwendigkeit Systeme doppelt vorzusehen, so daß Gerät zur visuellen Sprachwiedergabe nur lokal am Kommunikationsort benötigt wird und sich damit die vorliegende Erfindung vergleichsweise kostengünstig realisieren läßt. Außerdem sind die erfindungsgemäßen Vorrichtungen und Systeme benutzerfreundlich, indem sie den Benutzer systematisch durch das Verfahren zum Einsatz und zur Steuerung des Arbeitsgangs zur visuellen Sprachdarstellung führen.

Ein weiteres wichtiges Merkmal der erfindungsgemäßen Vorrichtung, des Verfahrens und der Systeme besteht darin, daß es sich um ein offenes System handelt, so daß jeder Benutzer des Geräts zur visuellen Sprachdarstellung mit jedem anderen kommunizieren kann, der Zugang zu einem Telefon hat, wodurch ein Kommunikationsmodell geschaffen wird, bei dem jeder mit allen kommunizieren kann, da ein modernes Telefon überall anzutreffen ist. Dieser Vorteil steht in deutlichem Kontrast zu den geschlossenen Systemen nach dem Stand der Technik, bei denen speziell nur für diese Zwecke ausgebildete Systeme an allen Kommunikationspunkten vorhanden sein müssen, wodurch ein Kommunikationsmodell entsteht, bei dem einer nur mit jenen paar anderen kommunizieren kann, die zu diesen spezialisierten zweckgebundenen Geräten und Systemen Zugang haben. Gemäß der vorliegenden Erfindung kann jeder Hörbehinderte über ein normales Telekommunikationsnetz mit jedem anderen Teilnehmer kommunizieren, ohne daß an einem dieser entfernten Orte, an denen sich der andere Teilnehmer befindet, eine besondere Ausrüstung benötigt wird. Dieses Merkmal eines offenen Systems ist wirklich revolutionär und bisher einmalig, da es erstmals eine universelle Möglichkeit zur Kommunikation mit Hörbehinderten über ein ganz normales Telekommunikationsnetz bietet, das sich irgendwo auf der Welt befindet.

Patentansprüche

1. Vorrichtung zur visuellen Wiedergabe von Sprache, **dadurch gekennzeichnet**, daß sie folgendes aufweist:
eine Netzschnittstelle (**110**), die mit einem ersten Kommunikationskanal (**103**) zum Empfangen eines ersten Tonsignals zur Bildung eines Tonempfangssignals koppelbar ist;
einen Hochfrequenzmodulator (**270**) zur Umwandlung eines Bildausgangssignals im Basisband in ein Hochfrequenz-Bildausgangssignal auf einem zweiten Kommunikationskanal (**227**) zur Bildanzeige; und
eine Prozessorengruppe (**130**), welche mit der Netzschnittstelle (**110**) und dem Hochfrequenzmodulator (**270**) gekoppelt ist und unter Ansteuerung durch einen Satz Programmbefehle in der Weise anspricht, daß sie das Tonempfangssignal in eine Sprachwiedergabe in Text:form umsetzt und weiterhin die Textdarstellung gesprochener Sprache in das Bildausgangssignal im Basisband umsetzt.

2. Vorrichtung nach Anspruch 1, dadurch gekennzeichnet, daß sie eine mit der Netzschnittstelle (**110**) und der Prozessorengruppe (**130**) gekoppelte Benutzerschnittstelle (**120**) zum Empfangen eines Steuersignals aus einer Vielzahl von Steuersignalen aufweist.

3. Vorrichtung nach Anspruch 2, dadurch gekennzeichnet, daß die Benutzerschnittstelle außerdem mit

einer physikalischen Schnittstelle für die Eingabe der Vielzahl von Steuersignalen koppelbar ist.

4. Vorrichtung nach Anspruch 3, dadurch gekennzeichnet, daß die physikalische Schnittstelle ein Telefon ist.

5. Vorrichtung nach Anspruch 3, dadurch gekennzeichnet, daß die physikalische Schnittstelle eine Tastatur ist.

6. Vorrichtung nach Anspruch 3, dadurch gekennzeichnet, daß die physikalische Schnittstelle ein Rechner ist.

7. Vorrichtung nach Anspruch 1, dadurch gekennzeichnet, daß in der Prozessoranordnung (**130**) eine Vielzahl von Betriebsarten vorgesehen ist, zu denen eine Telefonbetriebsart und eine Betriebsart mit visueller Sprachwiedergabe gehören, und daß die Prozessorengruppe (**130**) des weiteren mit Auswahl der Betriebsart mit visueller Sprachdarstellung auf ein Steuersignal anspricht.

8. Vorrichtung nach Anspruch 1, dadurch gekennzeichnet, daß die Prozessoranordnung folgendes umfaßt:
ein Mikroprozessor-Teilsystem (**260**);
einen mit dem Mikroprozessor-Teilsystem (**260**) gekoppelten Speicher; und
einen mit dem Mikroprozessor-Teilsystem (**260**) und dem Speicher gekoppelten Prozessor (**305**) zur visuellen Sprachwiedergabe.

9. Vorrichtung nach Anspruch 8, dadurch gekennzeichnet, daß der Prozessor (**305**) zur visuellen Sprachwiedergabe weiterhin folgendes umfaßt:
einen Prozessor (**307**) zur Spracherkennung; und
einen mit dem Prozessor (**307**) zur Spracherkennung gekoppelten Prozessor (**309**) für die Wiedergabe auf einem Bildschirm.

10. Vorrichtung nach Anspruch 8, dadurch gekennzeichnet, daß der Prozessor (**305**) zur visuellen Sprachwiedergabe weiterhin folgendes umfaßt:
einen Prozessor (**307**) zur Spracherkennung; und
einen mit dem Prozessor (**307**) zur Spracherkennung gekoppelten Untertitelkodierer (**311**).

Hierzu 4 Seite(n) Zeichnungen

Nummer: DE 197 50 439 A1
Int. Cl.[6]: A 61 F 11/04
Offenlegungstag: 20. Mai 1998

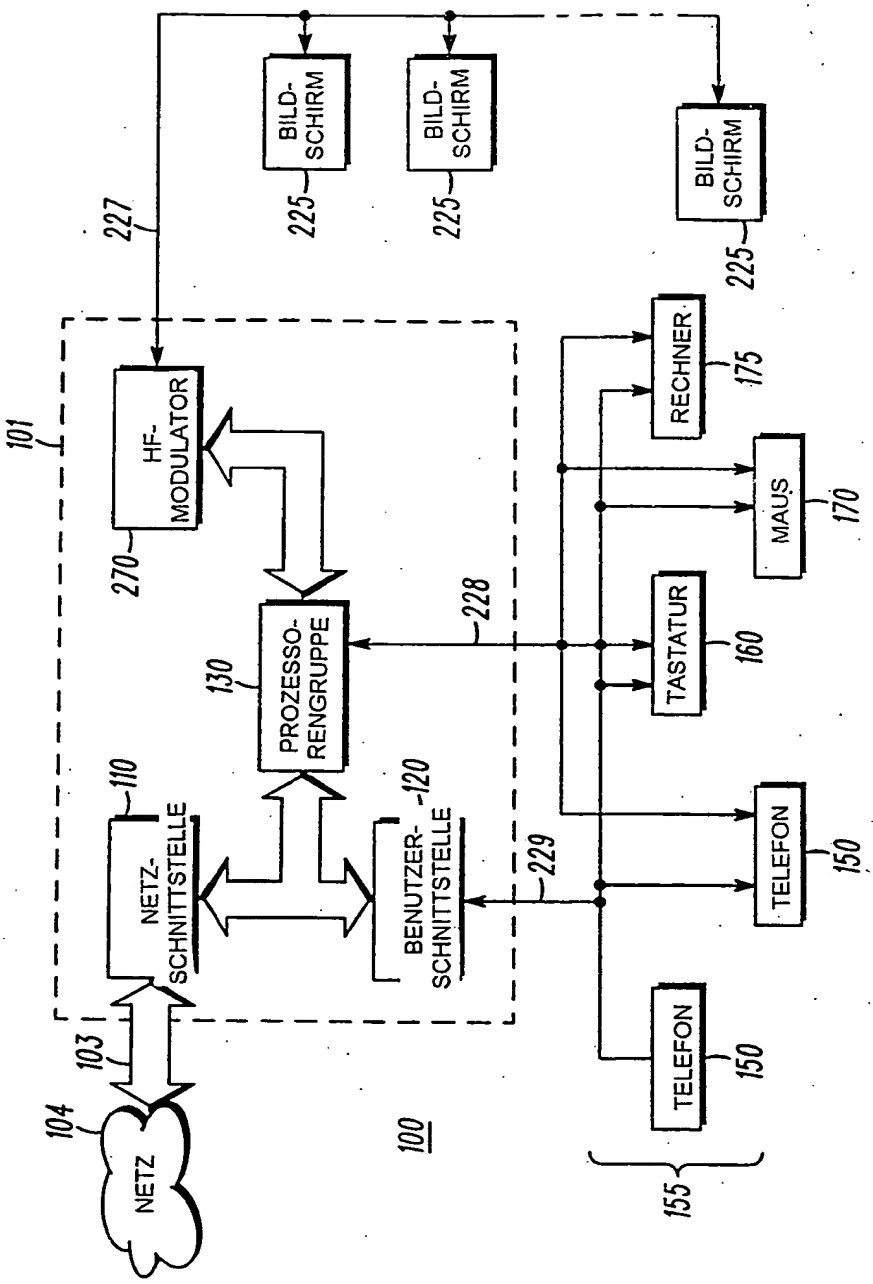


FIG.1

Nummer: DE 197 50 439 A1
Int. Cl.[6]: A 61 F 11/04
Offenlegungstag: 20. Mai 1998

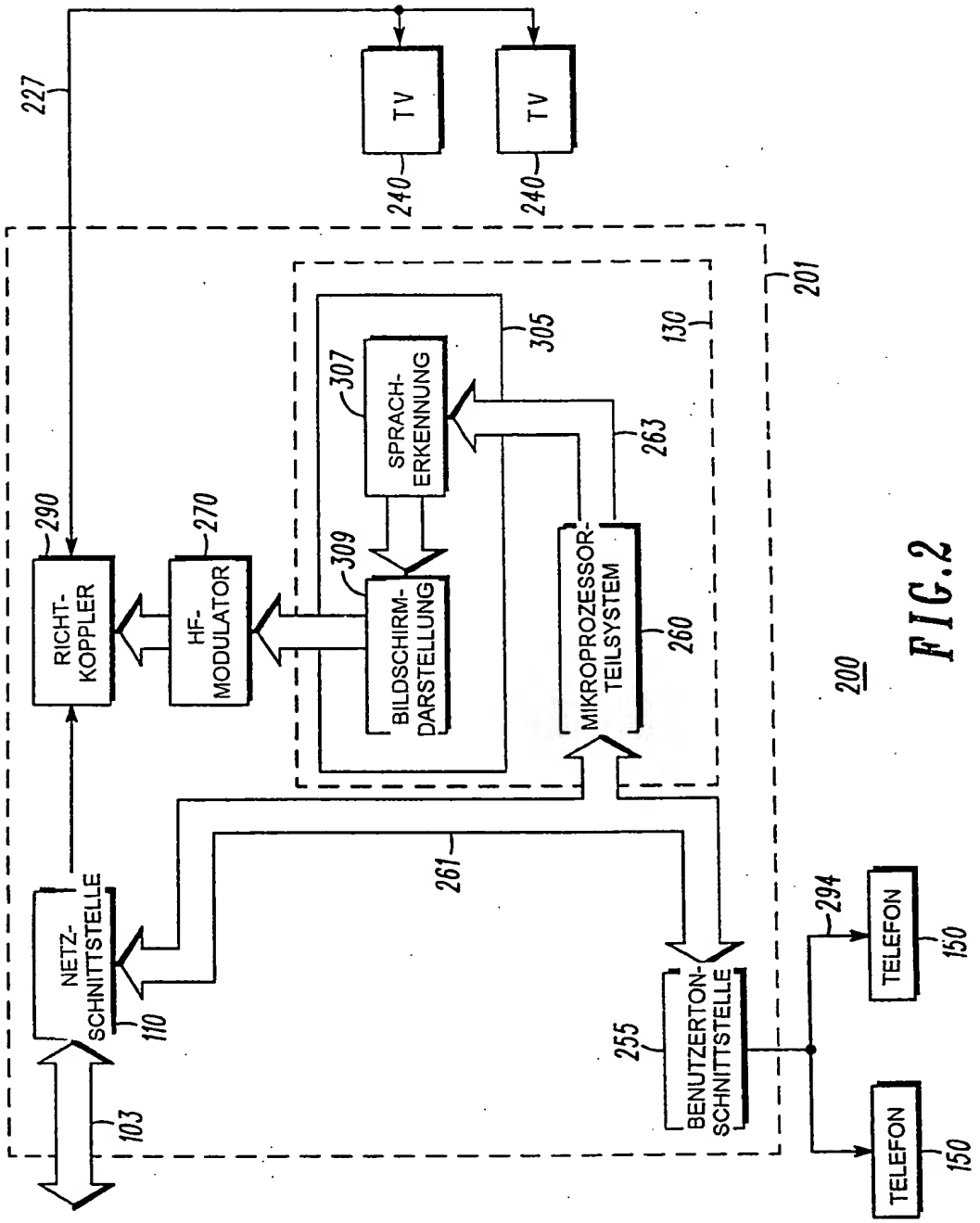


FIG.2

Nummer: DE 197 50 439 A1
Int. Cl.[6]: A 61 F 11/04
Offenlegungstag: 20. Mai 1998

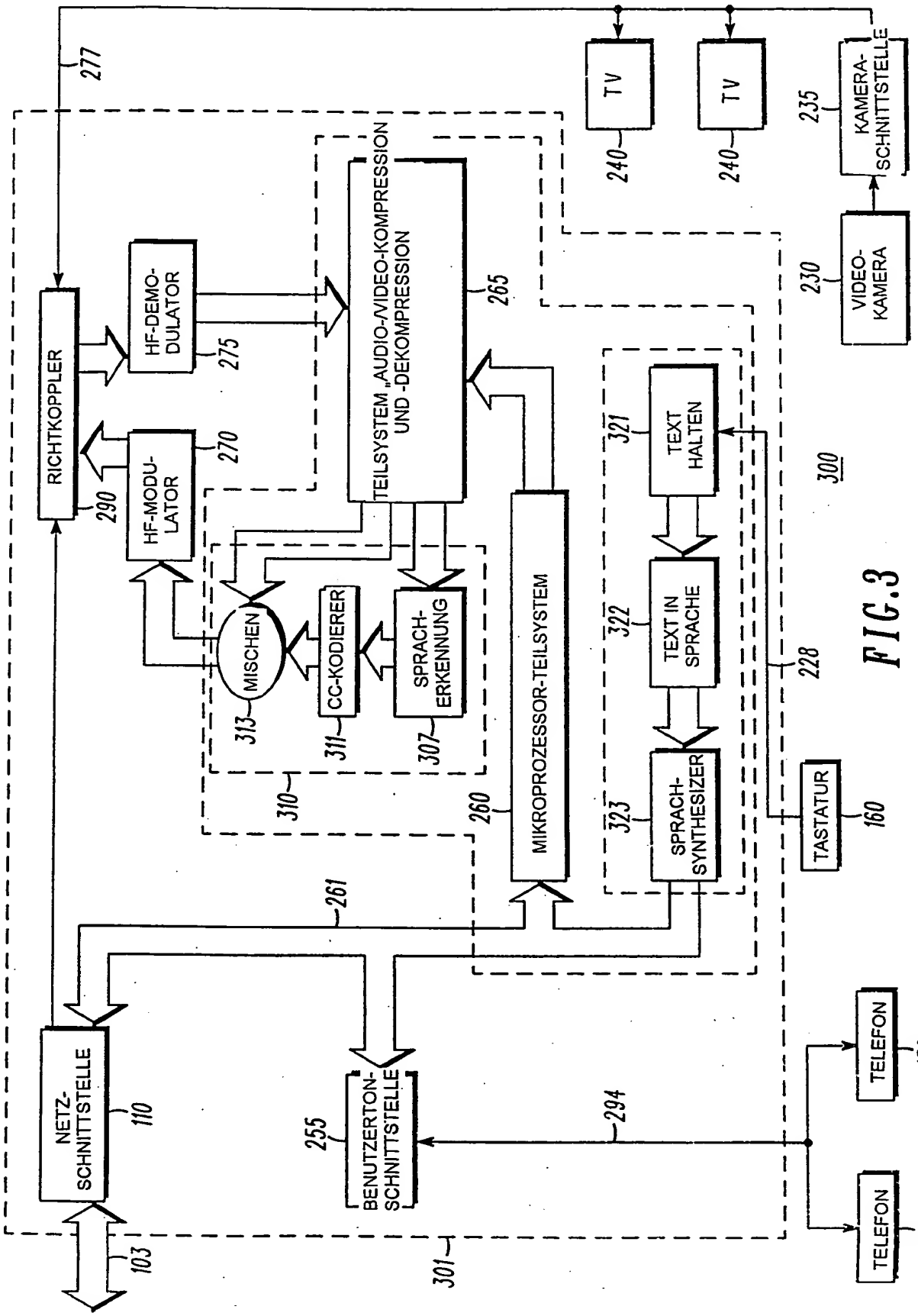
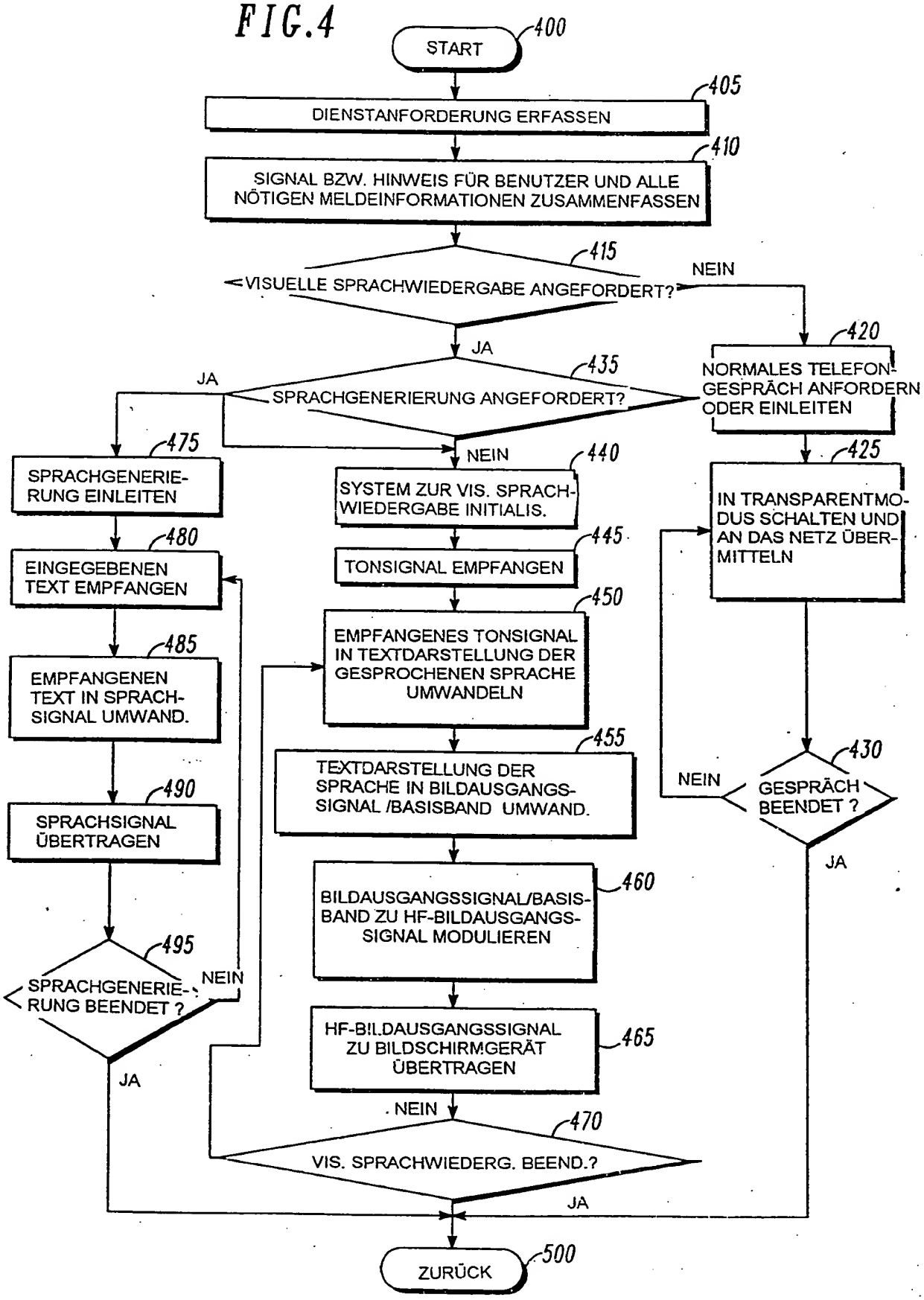


FIG.3

Nummer: DE 197 50 439 A1
Int. Cl.[6]: A 61 F 11/04
Offenlegungstag: 20. Mai 1998

FIG.4



BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☑ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☑ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.
As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.